Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation

Eirini P. Papapetrou[a,b], Mark J. Tomishima[c,d], Stuart M. Chambers[c], Yvonne Mica[e], Evan Reed[a,b], Jayanthi Menon[a,f], Viviane Tabar[a,f], Qianxing Mo[g], Lorenz Studer[a,c], and Michel Sadelain[a,b,1]

[a]Center for Cell Engineering, [b]Molecular Pharmacology and Chemistry and [c]Developmental Biology Programs, [d]SKI Stem Cell Research Facility, [e]Gerstner Sloan-Kettering Graduate School, and Departments of [f]Neurosurgery and [g]Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, 10065

Communicated by Mark Ptashne, Memorial Sloan–Kettering Cancer Center, New York, NY, May 1, 2009 (received for review April 8, 2009)

Human-induced pluripotent stem cells (hiPSCs) are generated from somatic cells by ectopic expression of the 4 reprogramming factors (RFs) Oct-4, Sox2, Klf4, and c-Myc. To better define the stoichiometric requirements and dynamic expression patterns required for successful hiPSC induction, we generated 4 bicistronic lentiviral vectors encoding the 4 RFs co-expressed with discernable fluorescent proteins. Using this system, we define the optimal stoichiometry of RF expression to be highly sensitive to Oct4 dosage, and we demonstrate the impact that variations in the relative ratios of RF expression exert on the efficiency of hiPSC induction. Monitoring of expression of each individual RF in single cells during the course of reprogramming revealed that vector silencing follows acquisition of pluripotent cell markers. Pronounced lentiviral vector silencing was a characteristic of successfully reprogrammed hiPSC clones, but lack of complete silencing did not hinder hiPSC induction, maintenance, or directed differentiation. The vector system described here presents a powerful tool for mechanistic studies of reprogramming and the optimization of hiPSC generation.

fluorescent proteins | lentiviral vectors | silencing | stoichiometry

Reprogramming of human fibroblasts to a pluripotent embryonic stem cell (ESC)-like state has recently been achieved through retroviral-mediated gene transfer of the 4 transcription factors Oct-4, Sox-2, Klf-4, and c-Myc (1–3). This combination of factors emerged from an initial screen in mouse fibroblasts based on co-transduction of 24 candidate genes (4). The generation of iPSCs with this method has now been reported from mouse and human somatic cells using various types of vectors, including gamma-retroviral, constitutive, or doxycycline (DOX)-inducible lentiviral and adenoviral vectors, as well as plasmid and transposon/transposase transfection systems (1, 2, 5–12). Nonetheless, the stoichiometric and temporal requirements of factor expression during hiPSC induction, maintenance, and differentiation remain poorly defined.

Direct reprogramming is a slow and inefficient process, with estimated efficiencies in human cells ranging from 0.02% to 0.002% (1, 2, 5). Many on-going efforts aim to identify genetic or chemical factors that enhance iPSC generation (13–16). A problem faced by these investigations is the lack of a consistent way of reporting reprogramming efficiency, since effects on factor delivery cannot be separated from genuine effects on reprogramming efficiency. Furthermore, the lack of proper assessment of factor delivery and expression obscures the comparison of reprogramming frequencies across different studies. The low efficiency of direct reprogramming may, at least in part, be accounted for by the requirement for a stringent stoichiometry of reprogramming factor expression permissive for successful reprogramming.

Several studies have reported silencing of the vector-encoded reprogramming factors in iPSCs generated with gamma-retroviral vectors (1, 17, 18). The silencing of lentiviral vectors in iPSCs is less known (2, 19), and therefore their suitability for reprogramming purposes remains controversial (20). Furthermore, the significance of vector silencing for reprogramming and its impact on differentiation is unclear. Although it has been proposed that silencing of factor expression is required for successful reprogramming and multilineage differentiation (3, 21, 22), it remains unclear whether vector silencing constitutes a requirement or an epiphenomenon of the reprogramming process.

Here we present a vector system for iPSC induction that allows for simultaneous real-time tracking of expression of the 4 individual transgenes in single cells during hiPSC induction, maintenance and directed differentiation. Using this system, we show that expression of the 4 RFs at an optimal stoichiometry is critical for efficient reprogramming, and we study in detail the kinetics of silencing or the 4 transgenes during the reprogramming process.

Results

Reprogramming of Human Fibroblasts with Bicistronic Vectors Coexpressing Each Reprogramming Factor with a Distinct Fluorescent Protein. We constructed lentiviral vectors that co-express each of the 4 RFs, Oct4, Sox2, Klf4, and c-Myc together with a fluorescent protein linked by a 2A peptide (Fig. 1*A*). We selected a combination of 4 fluorescent proteins, vexGFP (violet light excited-green fluorescent protein), mCitrine, mCherry, and mCerulean (23, 24), which can be separated by fluorescence microscopy and flow cytometry, thus enabling parallel monitoring of each reprogramming factor expression in real time (Fig. 1*B* and see below). Immunoblots of cells transduced with these vectors revealed correct processing of the 2A-linked gene products (Fig. S1 *A–D*). All 4 vectors were found to yield similar titers of 1.5–2 $\times$ 10^6 TU/mL. By co-transduction of human fetal fibroblasts (MRC-5) with these 4 vectors hiPSC lines were generated that expressed pluripotency markers (Fig. S2*A*) and could be directed to differentiate into derivatives of all 3 germ layers (Fig. S2*B*). Microarray-based global gene expression analysis revealed gene expression patterns highly similar to hESCs, but not to the parental human fibroblasts (Fig. S2*C* and Table S1). Furthermore, these hiPSC lines induced teratomas when s.c. injected into immunocompromised mice (Fig. 1*C*).

Requirements of Factor Expression for Efficient Reprogramming. We first took advantage of this traceable vector system to estimate what fraction of human fibroblasts co-expressing all 4 factors successfully reprogram (Fig. S2*D*). We co-transduced human fetal fibroblasts (MRC-5) with the 4 vectors at equal multiplicity of infection (MOI)

CELL BIOLOGY

Author contributions: E.P.P., M.J.T., L.S., and M.S. designed research; E.P.P., M.J.T., S.M.C., Y.M., E.R., J.M., and V.T. performed research; E.P.P., Q.M., and M.S. analyzed data; and E.P.P. and M.S. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: m-sadelain@ski.mskcc.org.

This article contains supporting information online at www.pnas.org/cgi/content/full/0904825106/DCSupplemental.

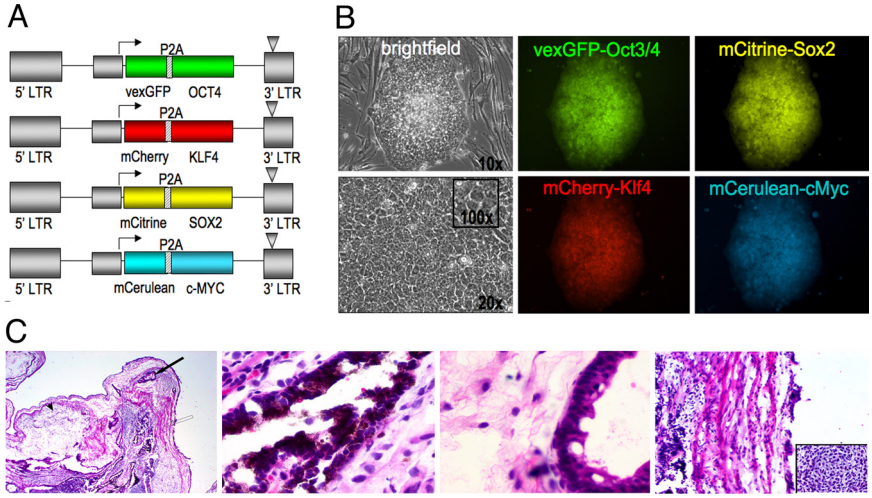

Fig. 1. hiPSCs derived from bicistronic vectors co-expressing each RF linked to a fluorescent protein. (*A*) Schematic representation of the vectors used in this study. LTR, long terminal repeat. (*B*) hiPSC colony at day 11 after transduction. (*C*) Hematoxylin and eosin staining of histological sections of a teratoma derived from line iPS-27. From *Left* to *Right*: Low power image demonstrating areas of heterogeneous differentiation: neuroectoderm (black arrow), smooth muscle (blank arrow), and primitive myxoid tissue (arrowhead), 4×. Pigmented epithelial tissue compatible with retinal neuroectoderm, 40×. Intestinal like epithelium including goblet cells (endoderm), 40×. Smooth muscle tissue, 20×. *Inset* demonstrates a high power image of immature mesenchymal tissue, potentially cartilage (40×).

and determined the percentage of quadruple transduced cells by flow cytometry 5 days after transduction. To calculate the frequency of fully reprogrammed cells, we performed immunostaining for Tra-1–81, as described in the *Methods*, and macroscopically enumerated positive colonies on day 20 after transduction. Thus, through direct quantification of quadruple transduced cells coupled with a clonal readout of hiPSC derivation, we were able to accurately estimate the reprogramming frequency to be 0.4–1% of de novo quadruple transduced fetal human fibroblasts (Table S2).

We then sought to investigate the effect of the absolute levels of expression of the 4 vector-encoded RFs on efficiency of direct reprogramming. Transduction of human fibroblasts at increasing MOI results in concomitant increase in the percentage of quadruple transduced cells, as well as in the level of expression (Fig. 2 *A* and *B*). Importantly, titration of the MOI of each vector results in linear titration of the mean fluorescence intensity (MFI) of the corresponding fluorescent protein (Fig. 2*B*). Our vector design also ensures that titrated expression of each fluorescent protein corresponds to titrated levels of expression of each linked RF (Fig. S1*E*). Increasing the levels of expression of the 4 transgenes simultaneously, as exemplified in Fig. 2*B*, had no effect in reprogramming efficiency of quadruple transduced cells. This result was corroborated by fluorescence-activated cell sorting (FACS) experiments that showed that the reprogramming efficiency of quadruple transduced human fibroblasts expressing all 4 transgenes at similarly high or low levels was the same (Fig. S3).

We therefore sought to investigate the effect that stoichiometric deviations in factor expression impose in the efficiency of reprogramming. We hypothesized that if differences in ratios of RF expression affect reprogramming efficiency in single cells, we could observe similar trends in bulk populations of quadruple transduced cells. Starting from a vector proportion of 1:1:1:1, we altered the stoichiometric ratio by varying the amount of 1 vector at a time while keeping the other 3 constant, as detailed in Fig. 2*C*. Flow cytometric analysis of factor expression 5 days after transduction was used to identify groups in which expression levels of any 3 factors (provided in constant amounts) maintained the same MFI and those were selected for comparison across permutations of expression of the fourth factor (Fig. S4 *A* and *B*). Reprogramming efficiency was calculated by enumeration of Tra-1–81+ colonies as above (Fig. S4*C*). We again did not observe correlation of reprogramming efficiency with the absolute expression levels of any of the 4 factors. In contrast, changes in the relative ratios of expression mediated significant effects on the efficiency of reprogramming. Increasing relative Oct4 expression resulted in enhanced reprogramming efficiency, whereas increasing the relative ratio of either Sox2, Klf4, or c-Myc consistently decreased efficiency of hiPSC colony generation by more than 5-fold (Fig. 2*C Upper*). In contrast, relative decrease of Sox2, Klf4 or c-Myc showed little effect, while relative decrease of Oct4 was detrimental (Fig. 2*C Lower*). These data demonstrate that a stoichiometry of equal parts of all 4 vectors is highly effective, resulting in reprogramming efficiency of de novo quadruple transduced fibroblasts similar to that observed in the maximally efficient secondary system (5, 7). Most deviations from this 1:1:1:1 stoichiometry have unfavorable effects to the efficiency of reprogramming, with the exception of a relative Oct4 increase. The amount of Oct4 expression has been shown to be critical in ES cells and its up- or down- regulation both drastically alter the pluripotent cell phenotype (25). We therefore examined in more detail the effect of relative Oct4 expression levels in direct reprogramming. Although a 3-fold relative increase was favorable, further increases were detrimental (Fig. 2*D*). Similar titrations of relative expression of the remaining 3 factors over a wider range did not reveal any additional effect than the trends shown in Fig. 2*C*. Based on these results, the optimal RF stoichiometry consists of a combination of equal amounts of Sox2, Klf4, and c-Myc with a 3-fold excess of Oct4.

Pronounced Transgene Silencing Is a Hallmark of Successful Reprogramming but Residual Factor Expression Does Not Hinder Early Lineage Specification. To study vector silencing, we established a panel of successfully reprogrammed clones on the basis of hESC-like morphology, Tra-1–81 expression and HLA-ABC downregulation to levels similar to hESCs ($n = 38$) and a panel of clones that formed colonies with non-hES morphology, did not express pluripotency markers and retained high HLA-ABC expression ($n = 30$). Ectopic expression of all 4 RFs was markedly lower in the panel of hiPSC lines than in lines that had not successfully undergone reprogramming to hiPSC state (Figs. 3 and S5*A*).

Despite showing pronounced vector silencing, a number of hiPSC lines were found not to have completely silenced factor expression. Some hiPSC lines maintained expression of 1 or more factors (at levels 20-fold lower than in transduced fibroblasts and in not fully reprogrammed clones). It has been suggested that some

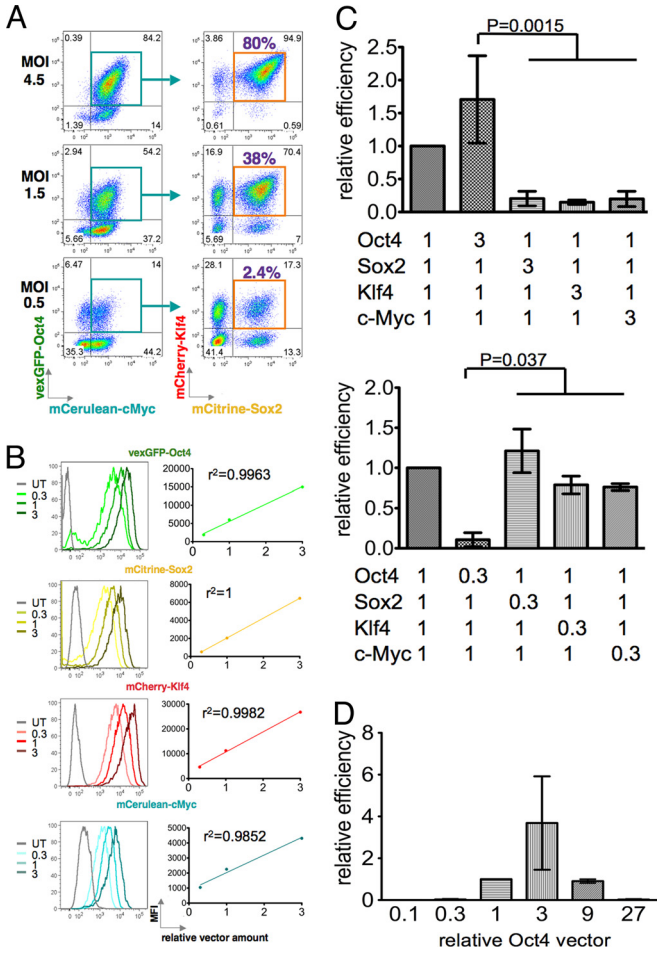

Fig. 2. Stoichiometric requirements of factor expression. (*A*) Representative flow cytometry analysis of MRC-5 fibroblasts on day 6 after transduction with the 4 vectors at increasing MOI as indicated. Numbers within plots denote percentage of cells in the respective quadrants. Dot plots in the *Right* are gated on vexGFP-Oct4+/mCerulean-cMyc+ double positive cells in the corresponding *Left*. The orange squares in the *Right* and the numbers above denote quadruple positive cells. (*B*) Representative flow cytometry analysis of MRC-5 fibroblasts on day 6 after transduction with titrated relative vector amounts (0.3, 1, 3), as indicated. UT, untransduced. (*C*) MRC-5 fibroblasts were transduced with relative vector amounts shown below the bars. Bars, mean ± SEM from 5 independent experiments. Absolute efficiency was calculated as the number of Tra-1–81+ colonies per number of plated cells divided by the fraction of quadruple positive cells estimated by flow cytometry analysis (as exemplified in *A*) at day 5 after transduction. Relative efficiency was calculated by normalizing absolute efficiency to the reference group (vector dose 1:1:1:1) to facilitate comparison across independent experiments. (*D*) Effect of relative Oct4 vector amount (*x* axis) on reprogramming efficiency (*y* axis), calculated as detailed in *C*. Bars, mean ± SEM from 3 independent experiments.

iPSC clones may fail to differentiate as a consequence of incomplete vector silencing (2, 13, 20). To address the significance of the residual vector expression, we investigated whether continued expression of 1 or more factors interferes with the pluripotent state and differentiation potential. We selected a panel of hiPSC lines, shown in Figs. 4 and Fig. S5*B*, on the basis of residual expression of either none of the factors (lines 14 and 27), Oct4 alone (lines 72 and 60), Oct4 and Sox2 (lines 51, 59 and 71), Oct4 and Klf4 (lines 76 and 62), or Oct4, Sox2, and Klf4 (lines 49 and 50). All 11 lines could be induced to differentiate into cells expressing endodermal and mesodermal markers (Fig. 4). All 7 out of 7 clones tested could also be induced to differentiate toward the neural lineage (Fig. 4). The ability of our system to monitor residual factor expression in single

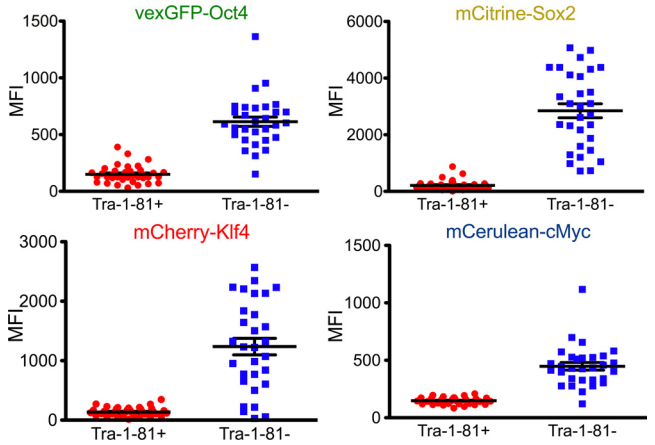

Fig. 3. Silencing of ectopic factor expression in hiPSC lines. MFI of each co-expressed fluorescent protein, in 38 reprogrammed (Tra-1–81+, red circles) and 30 non-reprogrammed (Tra-1–81-, blue squares) clones. Error bars, SEM.

cells revealed intraclonal variegation in most hiPSC lines (Figs. 4 and Fig. S5*B*). To test if residual factor expression biases the ability of individual cells to differentiate, we simultaneously monitored exogenous Sox2 and endoderm marker CXCR4 expression upon directed differentiation of hiPSC lines 50 and 59 (Fig. S5*C*). This analysis revealed lesser differentiation marker expression in the transgene-expressing fraction, suggesting that subsets of cells with continuous factor expression within a clone may differentiate less efficiently. These results taken together demonstrate that residual RF expression does not abolish the ability for early lineage specification toward any of the 3 germ layers, but may reduce the efficiency of directed differentiation.

To investigate whether changes in epigenetic status occurring during differentiation could cause re-activation or upregulation of RF expression, we closely monitored transgene expression during the course of differentiation. No induction of factor expression was observed in 10 out of 10 hiPSC lines during induction to endoderm, mesoderm, and neuroectoderm (Fig. S6*A*).

Vector Silencing Follows Acquisition of Pluripotent Cell Markers and Is Not Selective for the Reprogramming Factors. The observation that vector silencing is a characteristic of iPSC clones has led to the speculation that it constitutes a prerequisite for successful reprogramming (20, 21). However, silencing could alternatively be just a consequence of the acquisition of a pluripotent cell state, as retroviral vector silencing in human pluripotent cells is extensively documented (26, 27). To address this, we first investigated the kinetics of vector silencing in relation to epigenetic remodeling and reprogramming by monitoring expression of each fluorescent protein in quadruple transduced cells over the first 20 days of the reprogramming process. Tra-1–81 and HLA-ABC were used as markers for cells undergoing reprogramming, as HLA-ABC expression was found to be significantly lower in hESC and hiPSC lines than in fibroblasts. The expression levels of all 4 transgenes were uniformly diminished in Tra-1–81+ relative to HLA-ABC^{high} cells and found to gradually decrease over time (reaching 35–50% of initial levels of expression by day 20 and <5% by passage 5) (Figs. 5 *A* and *B* and Fig. S6*B*). Vector silencing with similar kinetics was also observed when a second pluripotency marker, SSEA3, was used. In contrast, transgene expression remained elevated in HLA-ABC^{high} cells. These results were corroborated by real-time imaging of selected hiPSC clones during early steps of the reprogramming process (Fig. S7). Further vector silencing was observed in established hiPSC lines in a time-dependent fashion, reaching a plateau after 12–15 passages (Fig. S6*B*). These results demonstrate that vector silencing is established over an extended period following

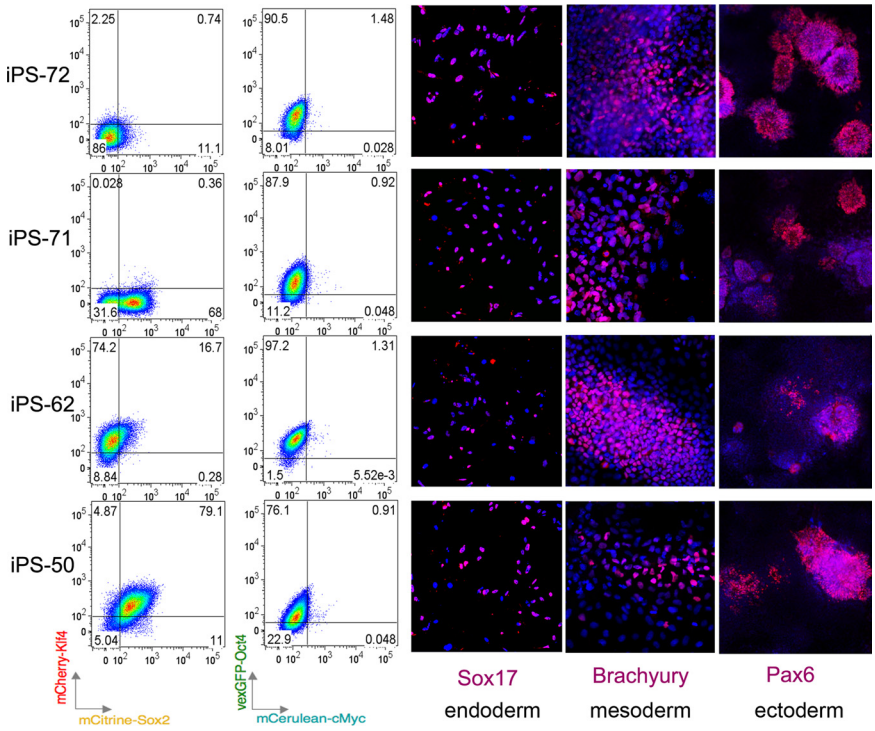

Fig. 4. Differentiation of hiPSC clones with incomplete RF silencing. *Left*, flow cytometry analysis of hiPS clones 72, 71, 62 and 50. *Right*, in vitro differentiation of the same clones in endoderm, mesoderm and ectoderm, followed by staining for Sox17, Brachyury, and Pax6, respectively.

acquisition of several criteria of pluripotency, such as morphology, growth, and pluripotency marker expression.

We further hypothesized that, if silencing were a requirement for reprogramming, it would be selective for the reprogramming vectors. To test this, we examined the state of silencing in fully reprogrammed (Tra-1–81+) versus partially reprogrammed (Tra-1–81-) clones generated from MRC-5 fibroblasts co-transduced with the 4 bicistronic vectors and an additional fifth

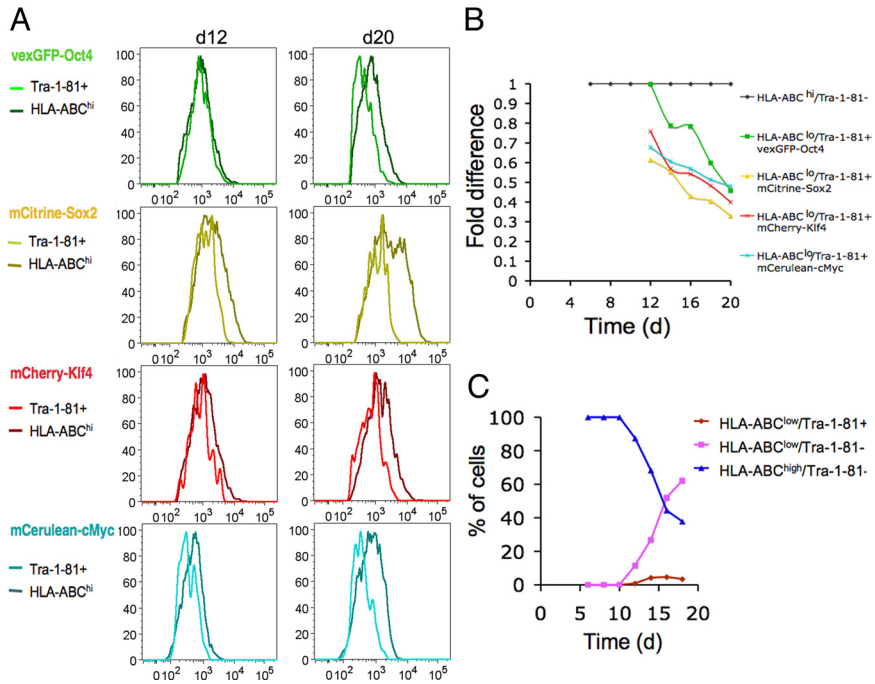

Fig. 5. Real-time simultaneous monitoring of vector-encoded RF expression and pluripotency marker induction. (*A*) Vector expression in Tra-1–81+ versus HLA-ABC^{high} cells on day 12 (*Left*) and 20 (*Right*) after transduction. (*B*) Vector silencing in Tra-1–81+ cells over time. (Day 0 denotes time of transduction.) Fold difference was estimated as ratio of MFI in the Tra-1–81+ cells to the MFI in HLA-ABC^{high} cells. (*C*) Appearance of pluripotency marker Tra-1–81 and downregulation of HLA-ABC during reprogramming. Shown is 1 of 3 independent time course experiments.

vector that has no impact on the reprogramming process, encoding dsRed-monomer alone. The silencing status of dsRed was found to be similar to the status of the 4 reprogramming vectors (Fig. S5*D*), establishing the absence of any specific selective pressure for RF silencing during reprogramming. The above data taken together suggest that silencing may be an epiphenomenon of the epigenetic state of pluripotent cells rather than a factor contributing to the acquisition of the pluripotent state.

Interestingly, the monitoring of marker expression in cells co-expressing all 4 factors revealed the emergence of a Tra-1–81-/HLA-ABC^{low} population concomitantly with the appearance of Tra-1–81+/HLA-ABC^{low} cells. The former population amplified markedly over time, rising to a much higher proportion than the latter fraction (Figs. 5*C* and Fig. S6*C*). These cells, which down-regulated HLA-ABC and gradually silenced vector expression similarly to hiPSCs, but failed to up-regulate pluripotency markers such as Tra-1–81 and SSEA3 (Fig. S6*C*), may represent intermediate or "abortive" products of the reprogramming process (13). Such cells indeed did not yield stable clones over time. Clones negative for Tra-1–81 that could be propagated (Fig. S5*A*) retained high HLA-ABC expression (at levels similar to the starting fibroblasts) and displayed a lack of vector silencing. This suggests that the Tra-1–81-/HLA-ABC^{low} cell population represents an intermediate state when cells have undergone some degree of epigenetic remodeling, but have not acquired self-renewal and/or pluripotency properties that enable their outgrowth under hiPSC culture conditions.

Discussion

Understanding the stoichiometric requirements and expression kinetics of the vector-encoded reprogramming factors is essential to standardize the generation of iPSCs and to better define the molecular events underlying the reprogramming process. Toward this goal, we have developed a vector system enabling accurate quantification of reprogramming efficiency and real-time monitoring of reprogramming factor expression.

Our results demonstrate in a direct prospective way that stoichiometry of reprogramming factor expression is a critical contributing factor to successful hiPSC induction. We establish that a stoichiometry of equal parts of all 4 vectors is highly effective, as most deviations from this stoichiometry have unfavorable effects to the efficiency of reprogramming, with the striking exception of a 3-fold relative Oct4 increase (Fig. 2 *C* and *D*). In view of this result, it is plausible that the initial screen by Takahashi and Yamanaka (4) selected for a combination of factors that mediate reprogramming in equal parts, introducing a selection bias, given that (*i*) all 4 vectors encoding the 4 RFs yield similar titers and (*ii*) successful iPSC generation requires co-transduction at high MOIs which tends to yield cells expressing similar relative levels of all 4 factors rather than a heterogeneous population expressing markedly different relative levels of the 4. The finding that increased relative levels of Sox2 are detrimental to the efficiency of reprogramming is consistent with the observation that inclusion of Sox2 in the RF mixture decreases iPSC generation from neural progenitor cells that express high levels of endogenous Sox2 (28, 29). The requirement for a specific factor stoichiometry is also consistent with the finding that the 4 RFs co-occupy promoters of many genes in iPSCs and ESCs and may co-operatively act during reprogramming (30). The stoichiometric requirements presented here will prove useful in guiding the implementation of optimized protocols for hiPSC induction and the rational design of improved reprogramming vectors, such as polycistronic and non-integrating vectors (9, 31, 32). Since our study on RF stoichiometry was limited to human fetal fibroblasts, it cannot be excluded that different cell types may have distinct stoichiometric requirements.

We further demonstrate that pronounced vector silencing is a hallmark of hiPSC clones, even when generated by constitutively expressed lentiviral vectors. This finding is consistent with the notion that successfully reprogrammed cells become independent of ectopic expression of the RFs and rely on endogenous factor expression (5, 20, 21, 27). Furthermore, our data show that vector silencing can serve as a criterion for selection of successfully reprogrammed iPSC clones. Our system greatly facilitates the rapid screening of hiPSC clones bearing silenced vectors, as well as the monitoring for potential reactivation of factor expression during differentiation processes in vitro or in vivo. The latter will be particularly useful in disease-specific hiPSC lines to separate potential phenotypes attributable to expression of the 4 transgenes rather than the underlying genetic background and, thus, facilitate the interpretation of studies aimed at deciphering pathogenetic mechanisms, studying the contribution of specific genetic lesions to the disease phenotype or therapeutic drug screening.

Recent studies on iPS clones generated with excisable vector systems demonstrate that residual exogenous factor expression does not hinder high-level chimeric contribution (11), although it perturbs the transcriptional profile (33) or the differentiation ability (10) of iPSCs. Our differentiation studies suggest that residual low-level factor expression does not prevent early lineage specification toward any of the 3 germ layers, but that it may subtly reduce the efficiency of differentiation.

Finally, the vector system presented here will greatly facilitate current efforts at optimization of direct reprogramming and screens for genetic or chemical factors that complement or substitute the 4 RFs. This system allows direct comparisons of reprogramming efficiency across different primary human cell types, treatments or other experimental variables in a controllable way allowing for normalization of outcomes based on efficiency of factor delivery and levels of expression. The reprogramming efficiency afforded by our system is comparable to that achieved in secondary fibroblasts bearing DOX-inducible vectors expressing the RFs (0.2–2%) (5, 7). The latter system provides the option of a genetically homogeneous background, albeit limiting the study of reprogramming to differentiated human iPSC-derived cells rather than primary cells and subject to instability of generated iPS clones upon doxycycline withdrawal (11, 17). Both systems remain susceptible to variegated expression due to position effects. Furthermore, the system presented herein can be used for packaging of integrase-deficient lentiviral vectors to facilitate tracking of transient reprogramming factor expression and screening of clones devoid of integrated vectors (8, 9, 34).

Methods

Plasmid Construction and Vector Production. Overlapping PCR was used to generate bicistronic expression cassettes encoding vexGFP, mCitrine, mCherry, and mCerulean linked by a P2A peptide preceded by a Gly-Ser-Gly linker (35) to the cDNAs of human Oct4, Sox2, Klf4 and c-Myc, respectively, which were cloned in a lentiviral vector under the transcriptional control of the human phosphoglycerate kinase (hPGK) promoter. Vector supernatants were produced as previously described (36).

hiPSC Generation. MRC-5 fibroblasts (ATCC) were transduced with supernatants of the 4 lentiviral vectors in the presence of 4 μg/mL polybrene for approximately 16 h. Five days after transduction, fibroblasts were plated at a density of 20,000 cells per 60-mm dish on a feeder layer of mitomycin C-treated mouse embryonic fibroblasts (MEFs) (GlobalStem). The next day, the medium was switched to hESC medium supplemented with 6 ng/mL FGF2 and was replaced every day thereafter. Twenty days after transduction, colonies with hESC morphology were mechanically dissociated and transferred into 24-well plates on MEFs. Cells were thereafter passaged with dispase and expanded to establish hiPSC lines.

For the time course experiments, MRC-5 fibroblasts were seeded on gelatin-coated 6-well plates at approximately 1×10^5 cells per well and transduced 24 h later. Forty-eight hours after transduction, the medium was replaced with hESC medium supplemented with 6 ng/mL FGF2 and 0.5 mM VPA.

Immunostaining. For immunofluorescence, the following antibodies were used: Nanog (AF1997, R&D), Tra1–60 and Tra-1–81 (MAB4360 and MAB4381, Chemicon), SSEA-3 and SSEA-4 (MC-631 and MC-813–70, DSHB). Appropriate

CELL BIOLOGY

secondary antibodies conjugated to Alexa488 or Alexa568 (Invitrogen) were used for detection. Nuclei were detected with Hoechst 33258.

For Tra-1–81 immunostaining, duplicate plates were fixed with 4% paraformaldehyde and incubated with an anti- Tra-1–81 antibody (Chemicon), followed by incubation with a horseradish peroxidase-linked anti-mouse IgM secondary antibody (Invitrogen). Staining of positive colonies was achieved by addition of chromogenic substrate 4-chloro-1-naphthol (Sigma). Macroscopic enumeration of positive colonies was done by 2 independent reviewers in a blinded fashion.

Gene Expression Profiling. Total RNA was isolated with RNeasy kit (QIAGEN). Samples were processed as independent triplicates. Whole genome gene expression analysis was performed on Illumina BeadArrays at the MSKCC microarray facility. Software R (http://www.r-project.org/) was used to perform all statistical computations. Moderated 2-sample *t*-test implemented in LIMMA package was used to examine whether genes were differentially expressed. Storey's *q*-value that controls positive false discovery rate (FDR) was used to correct for multiple testing for each contrast of interest. *Q*-values less than 0.05 were considered statistically significant.

Teratoma Formation. Undifferentiated hiPSCs were suspended at 1×10^7 cells/mL in DMEM containing 10% FBS. One million cells were injected s.c. in the flank of adult (3-month-old) NOD/SCID mice (Jackson Laboratory). Eight weeks later, the tumor was surgically dissected, fixed in 4% formaldehyde and embedded in OCT. Sections were cut at 7-μm thickness and stained with hematoxylin and eosin. All animal experiments were conducted in accordance with protocols approved by MSKCC Institutional Animal Care and Use Committee (IACUC) and following National Institutes of Health guidelines for animal welfare.

Flow Cytometry. Cells were dissociated with accutase, stained with Alexa Fluor (AF) 647-conjugated anti- Tra-1–81 and PE-Cy5-conjugated anti-HLA-ABC or AF 647-conjugated anti-CXCR4 antibodies (BD Biosciences) and analyzed in a LSRII cytometer (BD Biosciences). Analysis was performed with the FlowJo software (version 8.8.4; Tree Star). Cell sorting was performed on a MoFlo cell sorter (DakoCytomation).

In Vitro Differentiation. Endoderm and mesoderm were induced as previously described (37, 38). Briefly, for endoderm differentiation hiPSCs were passaged onto MEFs and expanded for 2–3 days before media was switched to endoderm induction medium (RPMI, 0.5% FBS, 2 mM L-glutamine, and 100 ng/mL activin A). For mesoderm differentiation hiPSCs were passaged onto MEFs and expanded for 6 days before media was switched to mesoderm induction medium (DMEM/F12, 10% FBS, 2 mM L-glutamine, 0.1 mM nonessential amino acids, and 0.1 mM beta-mercaptoethanol). In both protocols, the medium was replaced every other day and the cells were assayed on day 6. Neural induction was done as previously described (39).

Statistical Analysis. Linear regression analysis was performed using Prism software (version 5.0a; GraphPad).

ACKNOWLEDGMENTS. The authors would like to thank Kim Ping for excellent technical assistance, Ronald Ghossein for assistance in histological analysis, Christian Carson and Jurg Rohrer (BD Biosciences) and Jan Hendrikx for cell sorting. This work was supported by the Starr Foundation (Tri-Institutional Stem Cell Initiative, Tri-SCI-018) and New York State Stem Cell Science, NYSTEM (N08T-060).

1. Takahashi K, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131:861–872.
2. Yu J, et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318:1917–1920.
3. Park IH, et al. (2008) Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451:141–146.
4. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663–676.
5. Maherali N, et al. (2008) A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* 3:340–345.
6. Wernig M, et al. (2008) A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. *Nat Biotechnol* 26:916–924.
7. Hockemeyer D, et al. (2008) A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell* 3:346–353.
8. Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K (2008) Induced pluripotent stem cells generated without viral integration. *Science* 322:945–949.
9. Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S (2008) Generation of mouse induced pluripotent stem cells without viral vectors. *Science* 322:949–953.
10. Kaji K, et al. (2009) Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 458:771–775.
11. Woltjen K, et al. (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 458:766–770.
12. Yu J, et al. (2009) Human induced pluripotent stem cells free of vector and transgene sequences. *Science* 324:797–801.
13. Mikkelsen TS, et al. (2008) Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454:49–55.
14. Shi Y, et al. (2008) Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 3:568–574.
15. Zhao Y, et al. (2008) Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell* 3:475–479.
16. Huangfu D, et al. (2008) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotechnol* 26:795–797.
17. Brambrink T, et al. (2008) Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* 2:151–159.
18. Stadtfeld M, Maherali N, Breault DT, Hochedlinger K (2008) Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2:230–240.
19. Blelloch R, Venere M, Yen J, Ramalho-Santos M (2007) Generation of induced pluripotent stem cells in the absence of drug selection. *Cell Stem Cell* 1:245–247.
20. Maherali N, Hochedlinger K (2008) Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell* 3:595–605.
21. Jaenisch R, Young R (2008) Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132:567–582.
22. Wernig M, et al. (2007) In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448:318–324.
23. Anderson MT, et al. (1998) Pairs of violet-light-excited fluorochromes for flow cytometric analysis. *Cytometry* 33:435–444.
24. Shaner NC, et al. (2004) Improved monomeric red, orange, and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol* 22:1567–1572.
25. Niwa H, Miyazaki J, Smith AG (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24:372–376.
26. Teich NM, Weiss RA, Martin GR, Lowy DR (1977) Virus infection of murine teratocarcinoma stem cell lines. *Cell* 12:973–982.
27. Hotta A, Ellis J (2008) Retroviral vector silencing during iPS cell induction: An epigenetic beacon that signals distinct pluripotent states. *J Cell Biochem* 105:940–948.
28. Silva J, et al. (2008) Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol* 6:e253.
29. Eminli S, Utikal J, Arnold K, Jaenisch R, Hochedlinger K (2008) Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* 26:2467–2474.
30. Sridharan R, et al. (2009) Role of the murine reprogramming factors in the induction of pluripotency. *Cell* 136:364–377.
31. Sommer CA, et al. (2008) Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* 27:543–549.
32. Carey BW, et al. (2008) Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc Natl Acad Sci USA* 106:157–162.
33. Soldner F, et al. (2009) Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell* 136:964–977.
34. Mali P, et al. (2008) Improved efficiency and pace of generating induced pluripotent stem cells from human adult and fetal fibroblasts. *Stem Cells* 26:1998–2005.
35. Szymczak AL, et al. (2004) Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat Biotechnol* 22:589–594.
36. Papapetrou EP, Kovalovsky D, Beloeil L, Sant'angelo D, Sadelain M (2009) Harnessing endogenous miR-181a to segregate transgenic antigen receptor expression in developing versus post-thymic T cells in murine hematopoietic chimeras. *J Clin Invest* 119:157–168.
37. Watanabe K, et al. (2007) A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol* 25:681–686.
38. D'Amour KA, et al. (2005) Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat Biotechnol* 23:1534–1541.
39. Chambers SM, et al. (2009) Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* 27:275–280.